Expanding SPoRT RGBs and Machine Learning Techniques to Enhance Air Quality Monitoring in Southern Asia

Jonathan L. Case[1], Emily B. Berndt[2], Aaron R. Naeger[2], Robert A. Junod[3], Andrew T. White[3], Kevin K. Fuell[3], Connor H. Welch**[3]

[1]ENSCO, Inc.

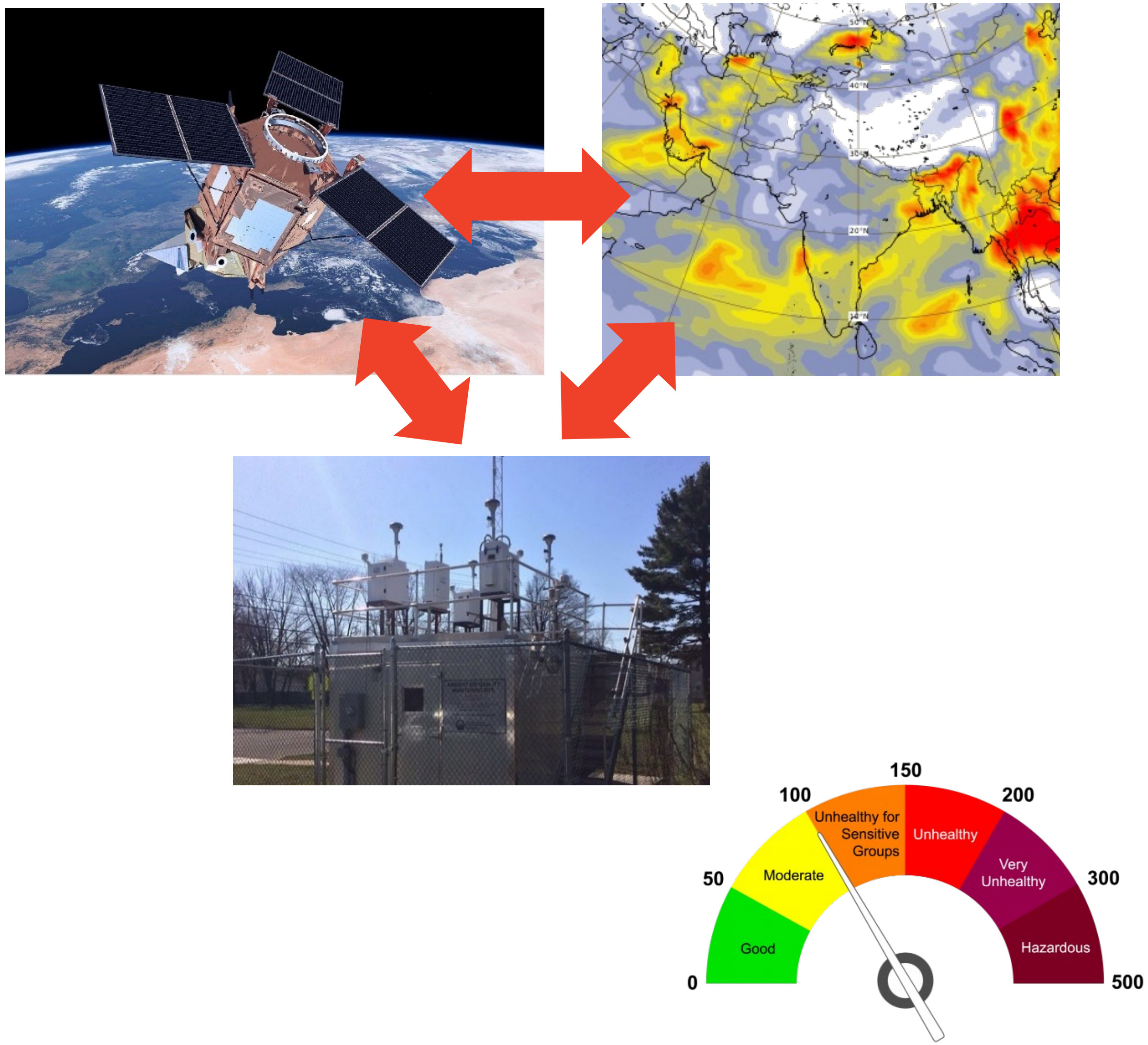
[2]NASA Short-term Prediction Research and Transition Center

[3]The University of Alabama in Huntsville

Air Quality Challenges in South Asia

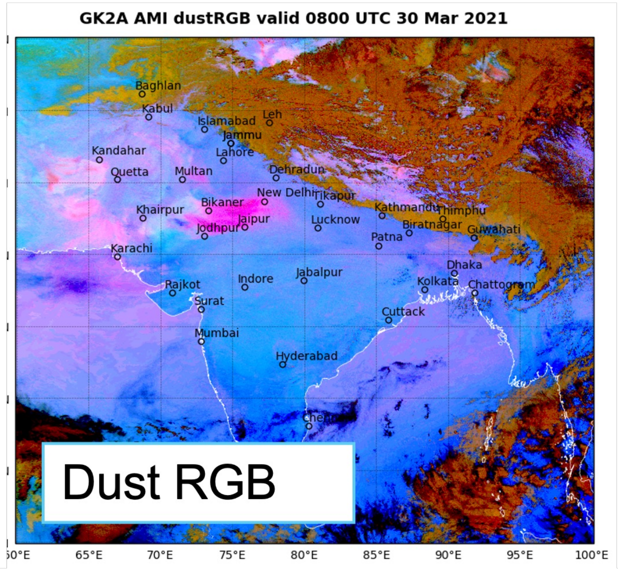
Short-term Prediction Research and Transition Center

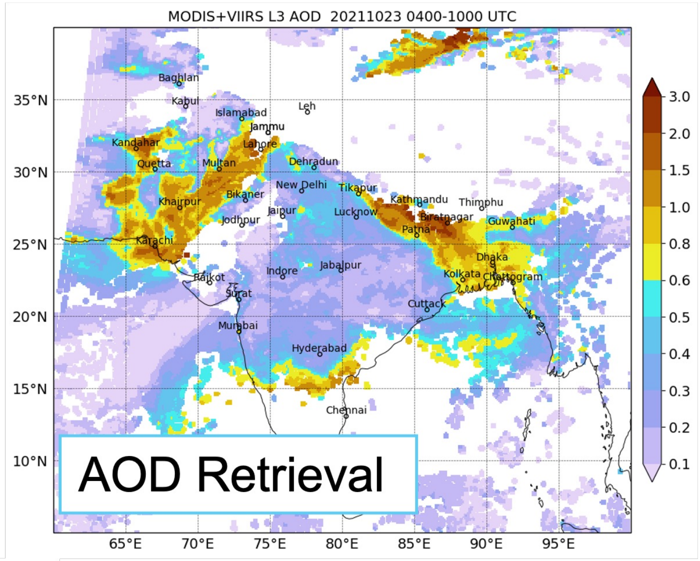
- Air pollution is a serious threat to human health
- Poor air quality common to region
- Difficult to monitor and predict due to strong / rapidly evolving emissions
- New generation satellite sensors can advance monitoring capabilities
- Ground-based sensors complement satellite & validate model data

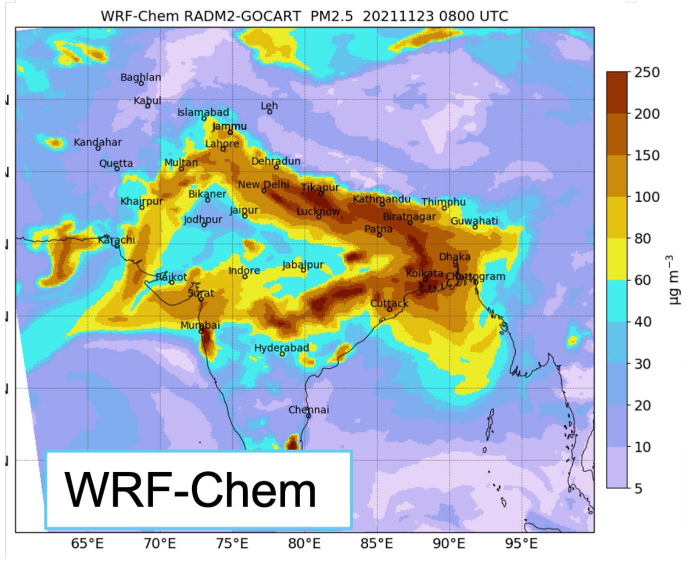


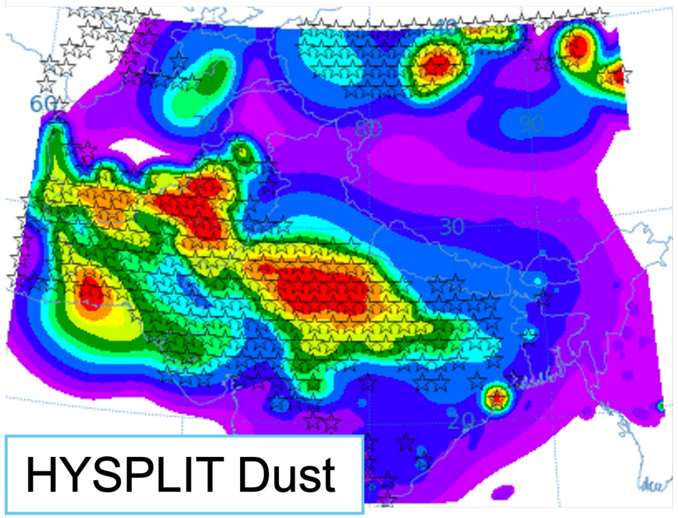
Foundational Products and Tools

- Red-Green-Blue (RGB) satellite products for monitoring dust, fog, fires, and smoke
- Trace gas / aerosol products to track pollution at surface and aloft
- Chemical transport model for predicting air quality and providing timely warnings
- Dispersion model for predicting dust concentrations & enabling rapid response to dust storms





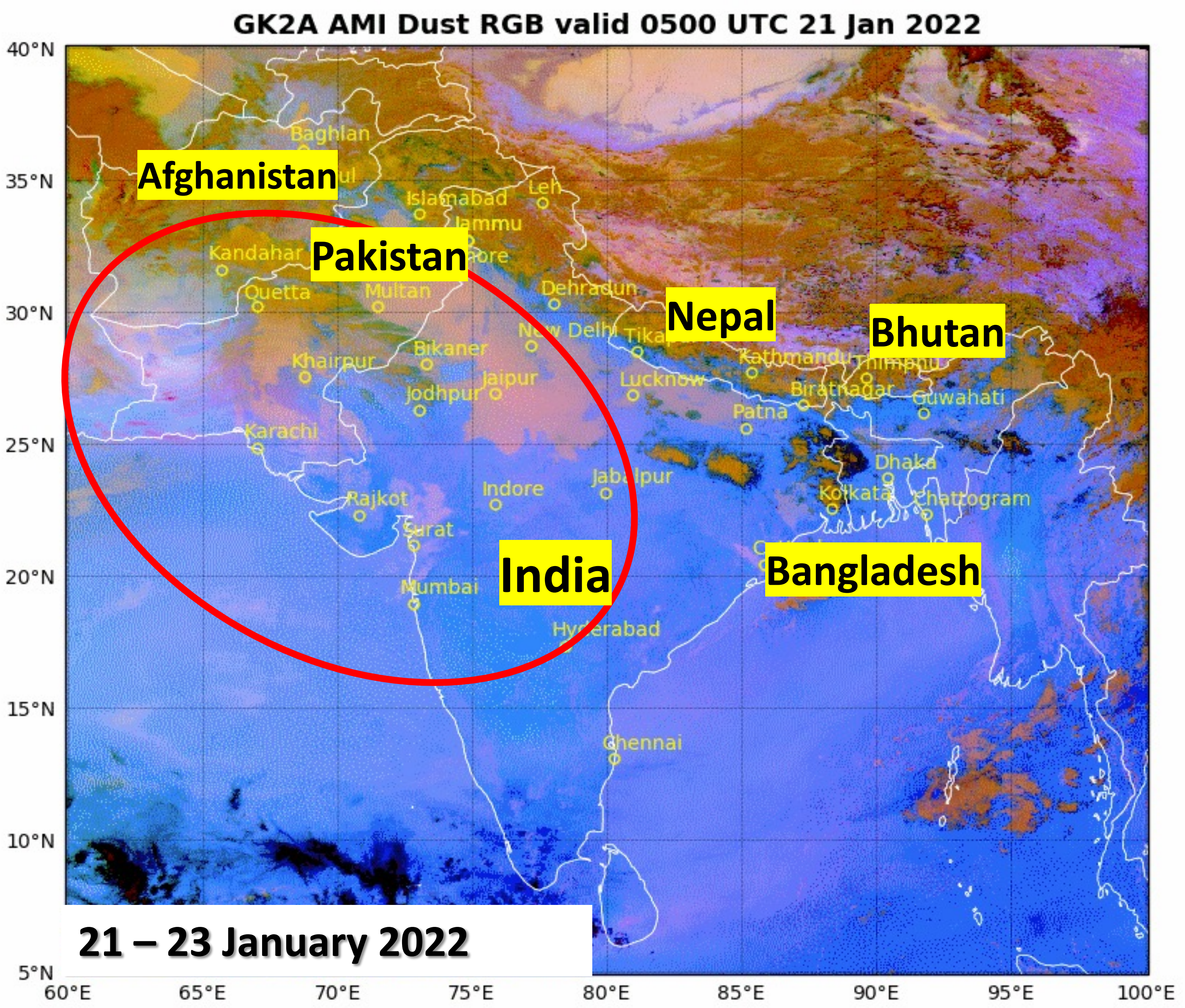




Satellite Products for Air Quality Monitoring

Satellite Product
○ Dust RGB
Application
○ Dust Plume Monitoring

Dust areas denoted by magenta shading (24-h coverage)

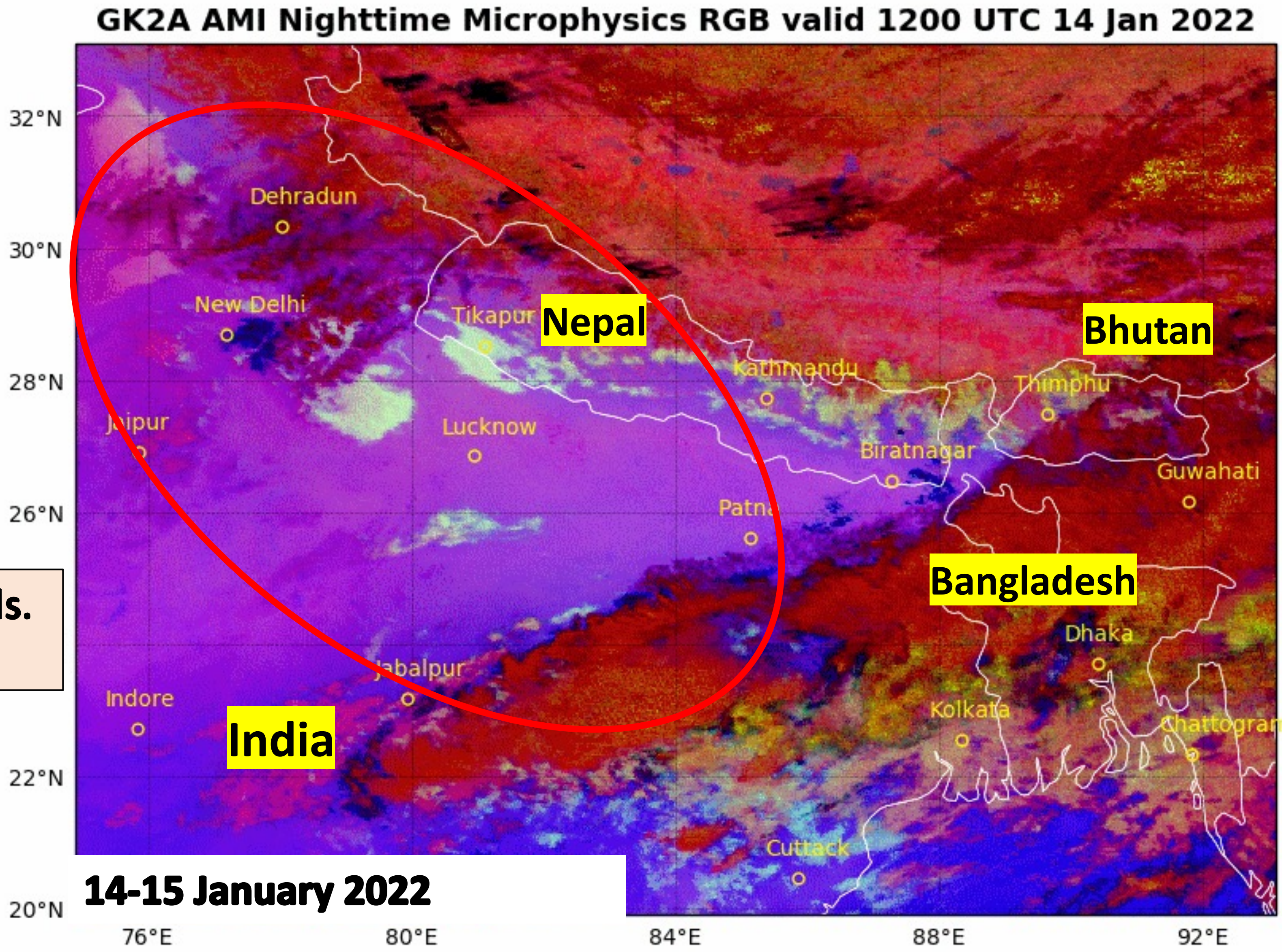


Satellite Products for Air Quality Monitoring

SPORT Short-term Prediction Research and Transition Center

Satellite Product
○ Nighttime Microphysics RGB
Application
○ Fog, smog, & low-cloud detection

Highlights fog & low clouds. (<u>Limitation</u>: night-only)

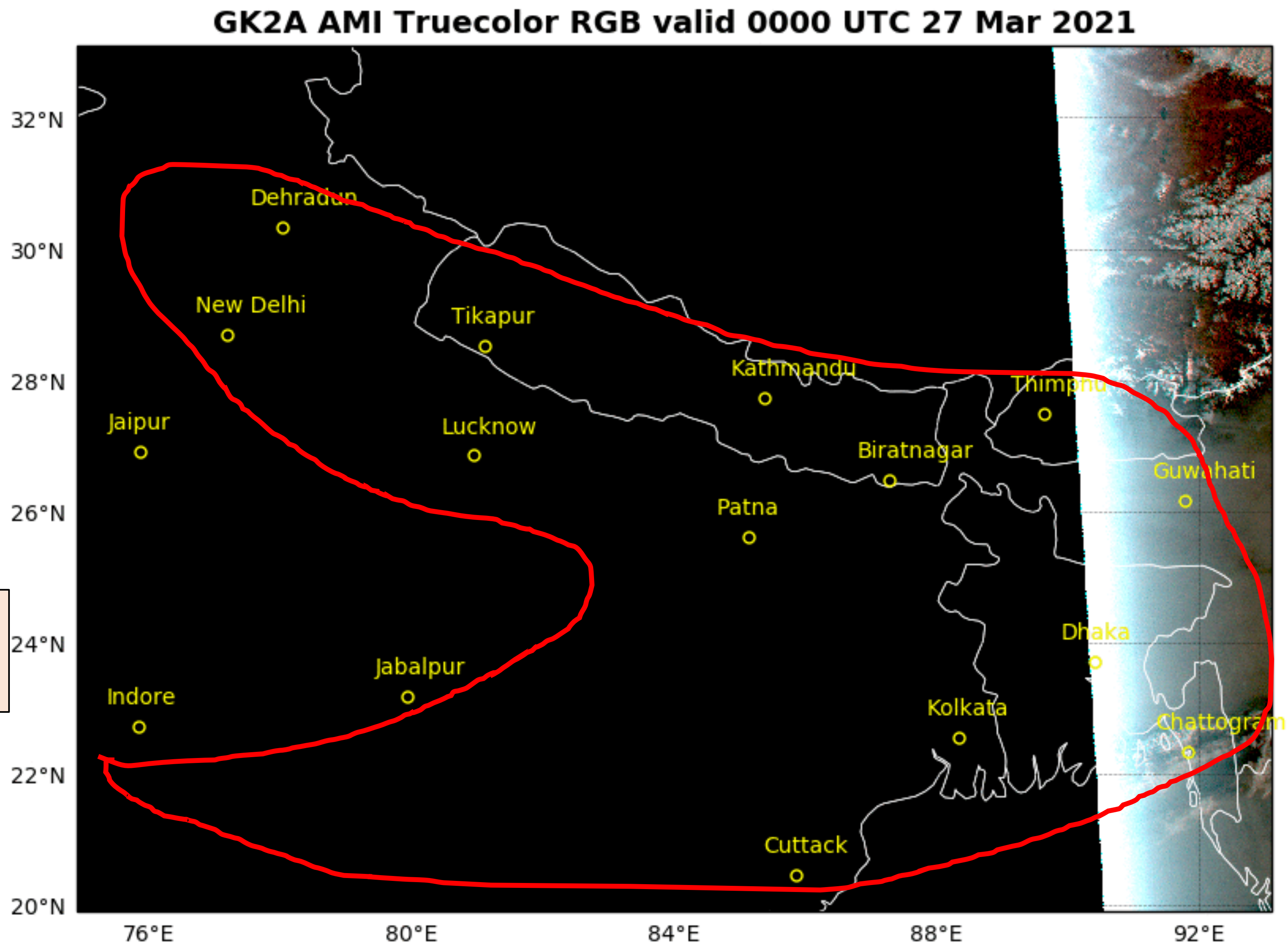


14-15 January 2022

Satellite Products for Air Quality Monitoring

Satellite Product
○ Truecolor RGB
Application
○ Land surface, clouds and smoke

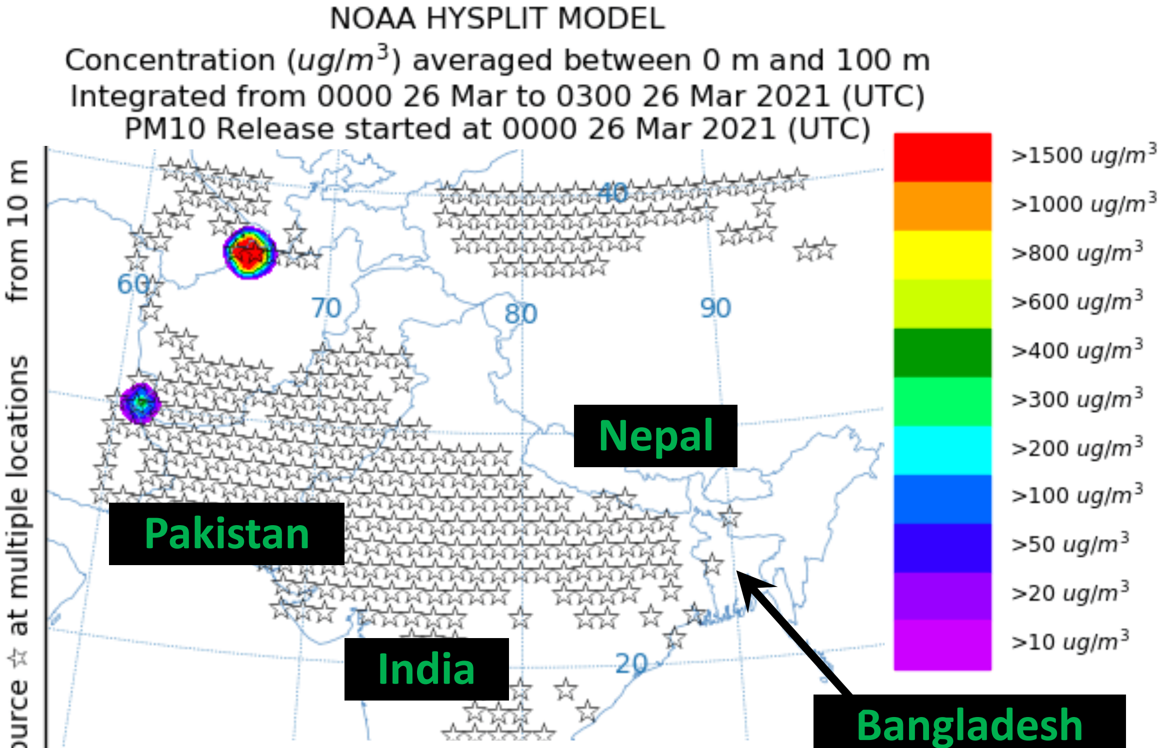
Truecolor shows regional smoke from biomass burning. (day only)



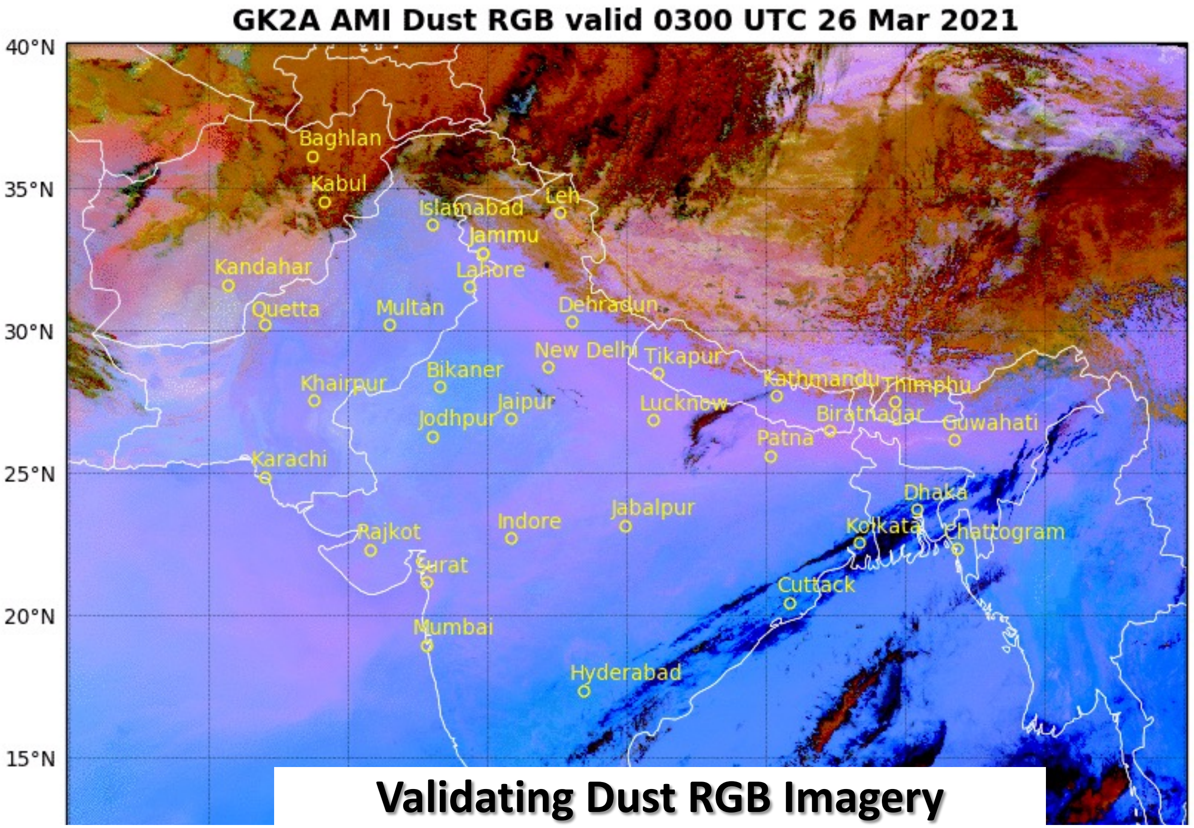
27 March 2021 Fire and smoke

Dust Dispersion Compared to Imagery

- 26-31 March 2021 dust event
  - HYSPLIT simple dust algorithm with monthly satellite emission climatology
  - <u>Room for improvement</u>:
    - ✓ Emission climatology from 2010 study (Draxler et al. 2010; *J. Geophys. Res.*)
    - ✓ Does not account for Land-use/Land-cover changes and/or seasonal anomalies

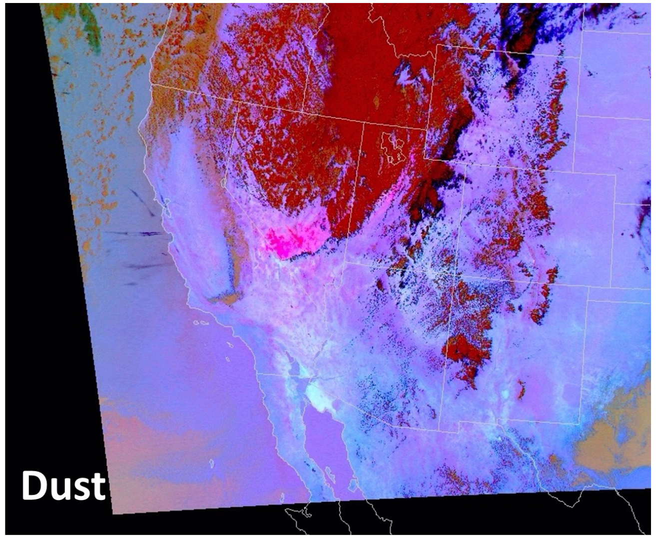


Simple Dust + Climo



Validating Dust RGB Imagery

Let's Improve Dust ID & Model Initialization

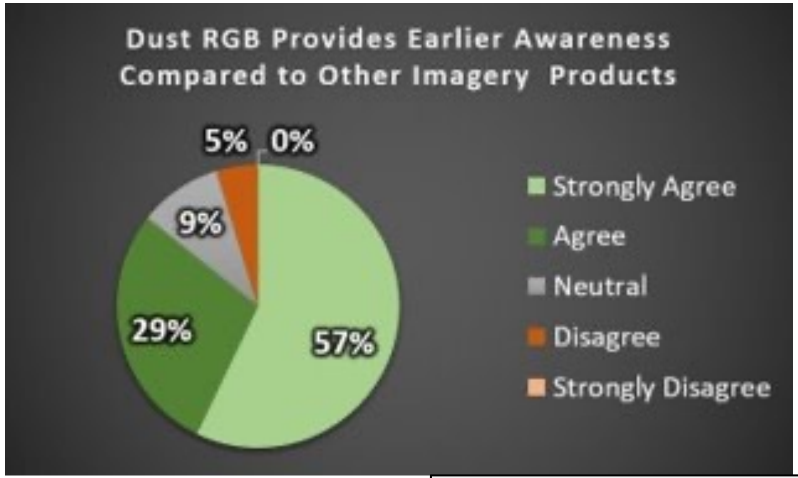


Early application & training of MODIS/ VIIRS RGBs to prepare for GOES-R

Early assessment of GOES-16 Dust RGB in operations

Limitations:
- Thin or dispersed dust
- Tracking long events into the night

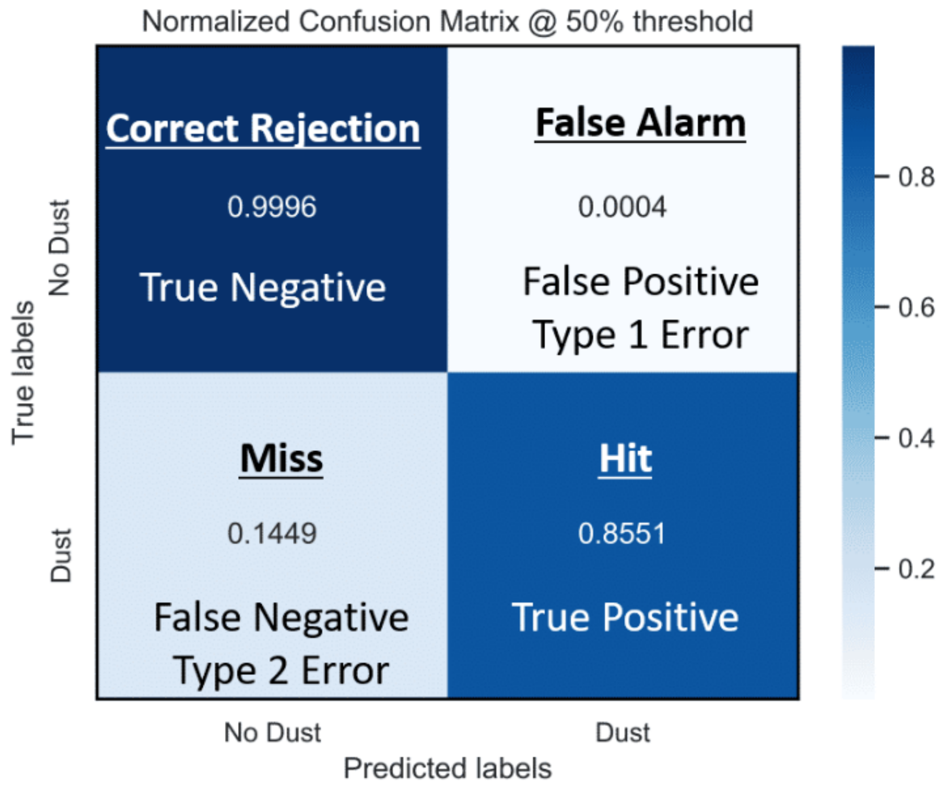
Use ML/AI to improve detection and interpretation



86% of Forecasters agree the Dust RGB provided earlier awareness of dust

DustTracker-AI Model Performance

Short-term Prediction Research and Transition Center



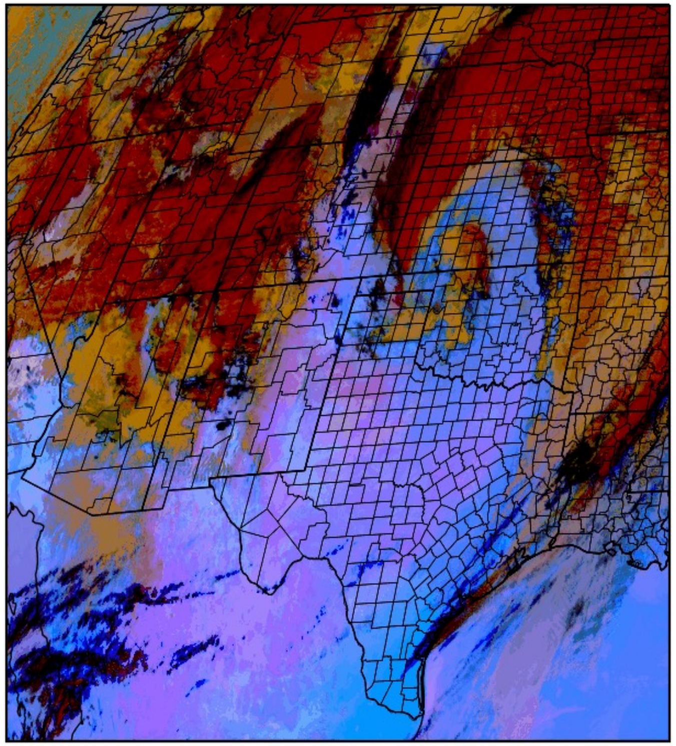
Finding Dust at Night

**Highly efficient at true 'dust' and 'no dust' classifications

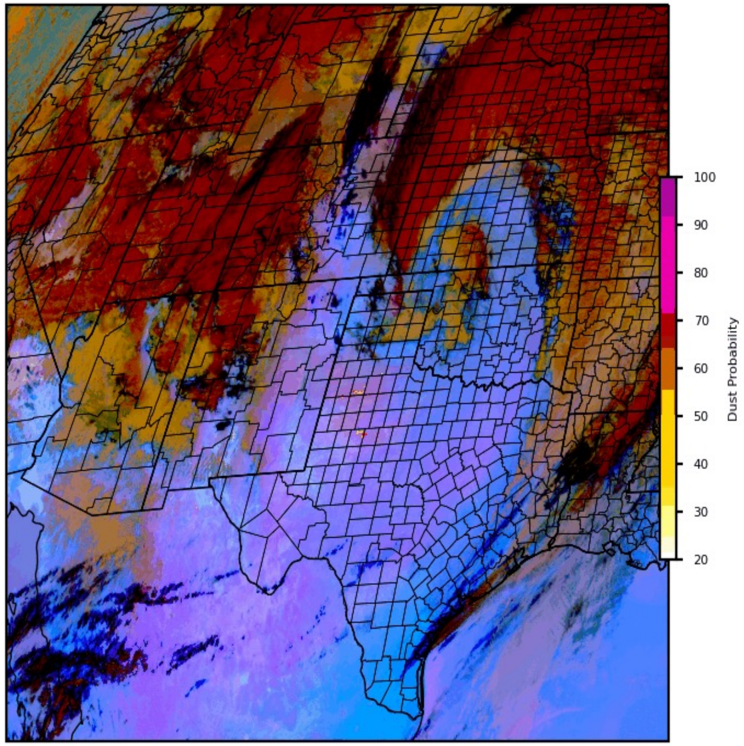
Recent Example: 14-15 Feb 2023

- **DustTracker-AI** has been developed as a Day/Night model to identify dust in complex night-time scenes and increase situational awareness
- End-User Assessment & Feedback on favorable and unfavorable cases will help expand the training database and improve the model

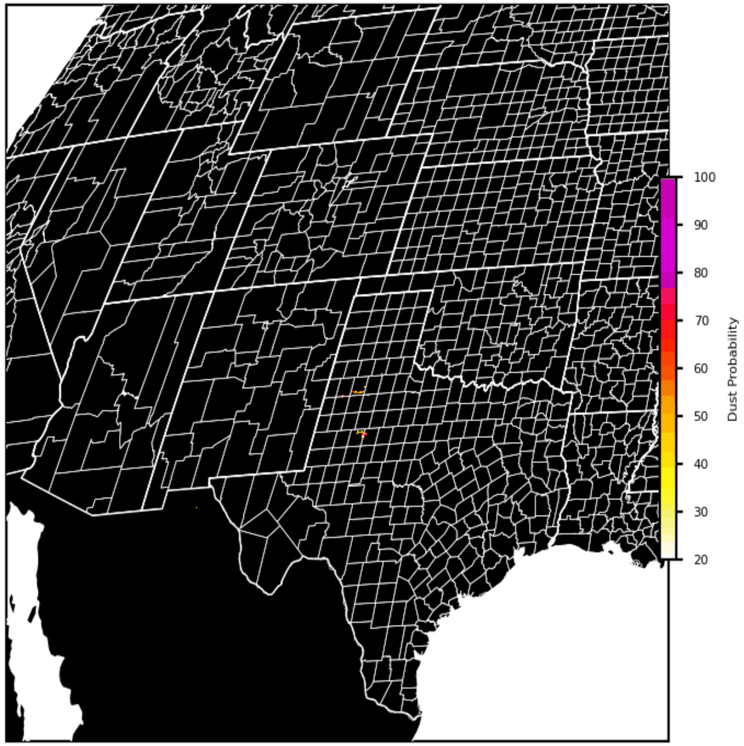
GOES-16 ABI Dust RGB 2023-02-14 17:51 UTC



GOES-16 ABI Dust RGB / Dust Probabilities 2023-02-14 17:51 UTC



Random Forest Dust Probabilities 2023-02-14 17:51UTC



Real-Time DustTracker-AI https://weather.ndc.nasa.gov/sport/viewer/?dataset=goeseastdustmodel&product=dust_prob

Merging AI and Physical Models:

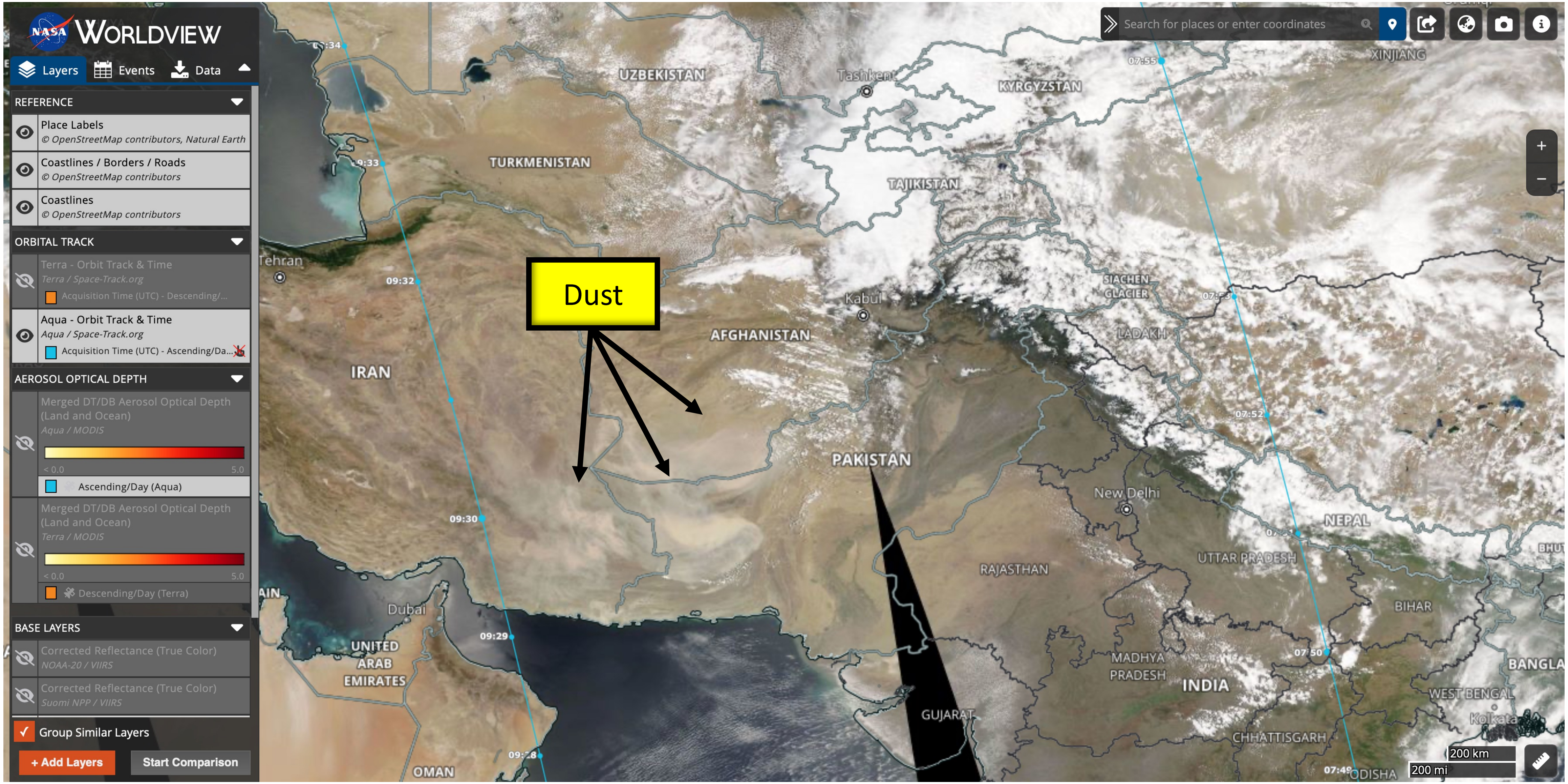
Short-term Prediction Research and Transition Center

Extend AI Models to Initialize Dust and Smoke Forecasts

- Develop AI models for dust and smoke identification over south Asia
  - **<u>Dust</u>**: Apply U.S. random forest model and train over south Asia
  - **<u>Smoke</u>**: Develop new AI model over south Asia (random forest? Deep learning?)
    - Could be more challenging! Satellite smoke ID a bigger challenge over region
    - Biomass burning: small but numerous; sub-pixel size (satellite ~1-2 km)
- Currently creating training data over south Asia for both phenomena
- Use AI Tracking models to inform dispersion models
  - Early hours of identification for emission source locations
  - Initialize dispersion model (i.e., HYSPLIT) with emission locations
  - Calibrate dispersion model parameters through validation against in situ and satellite observations

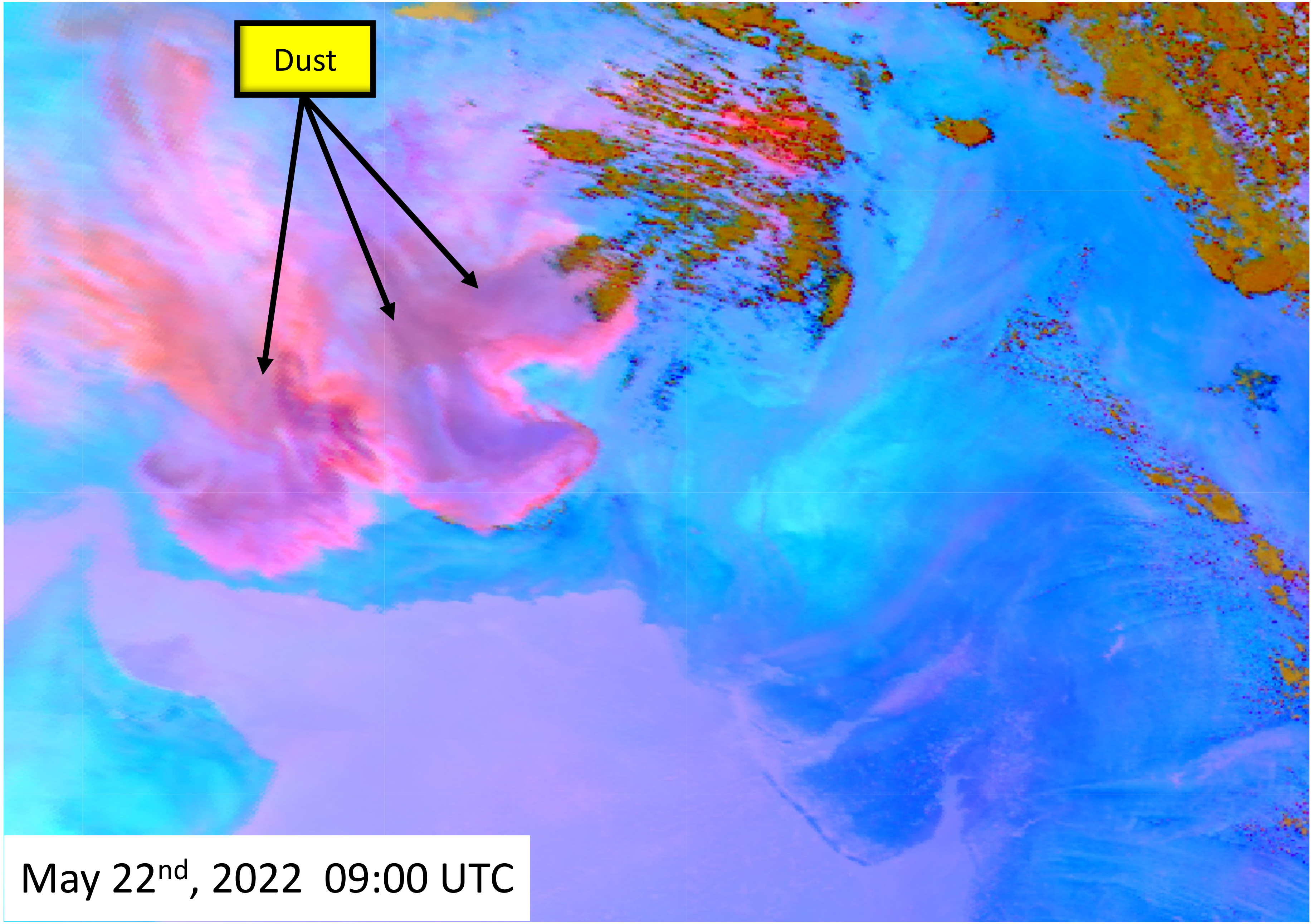
Example Training Data – Dust, Truecolor Aqua

May 22nd, 2022 09:30 UTC

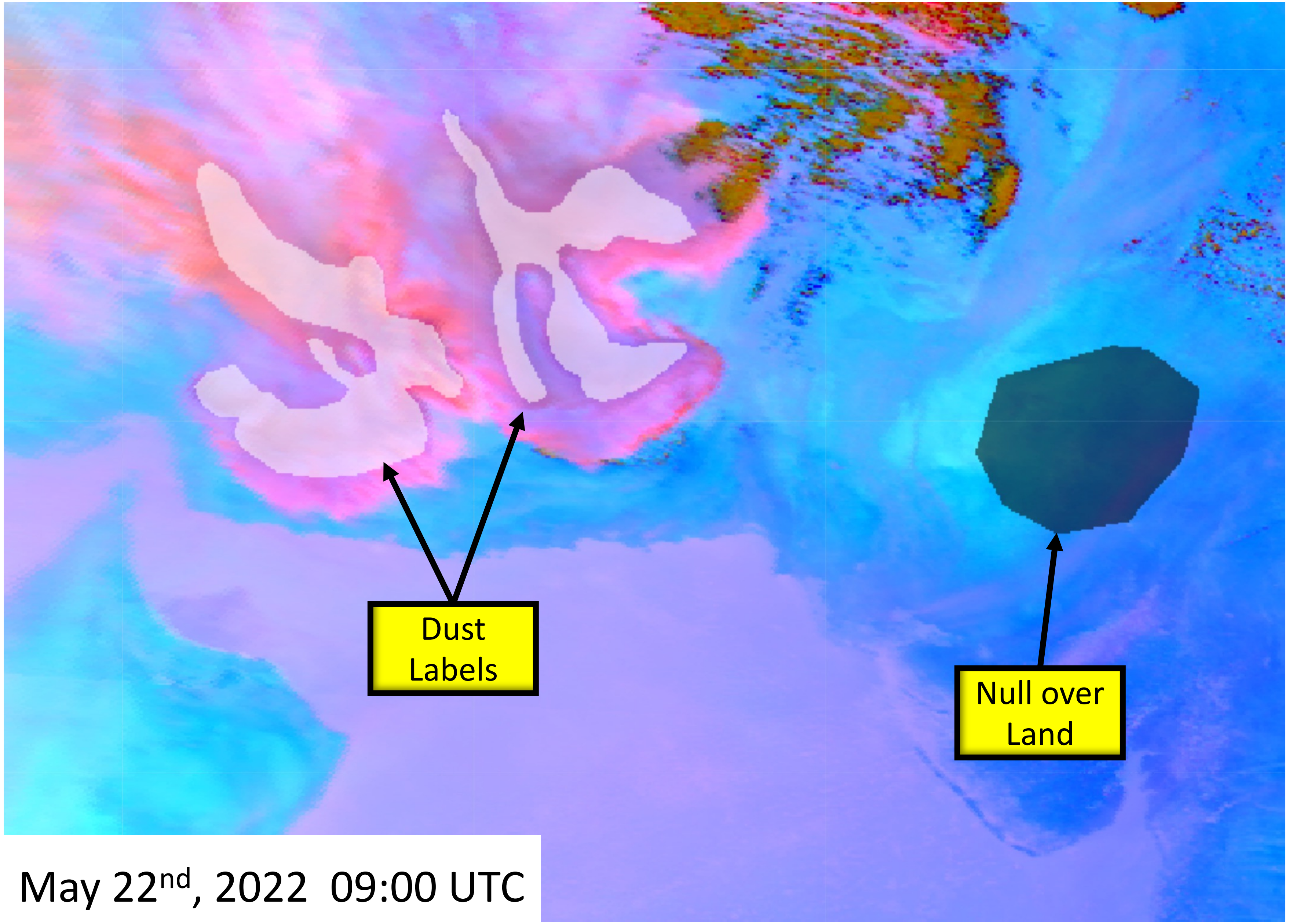


Example Training Data – Dust RGB

SPORT Short-term Prediction Research and Transition Center

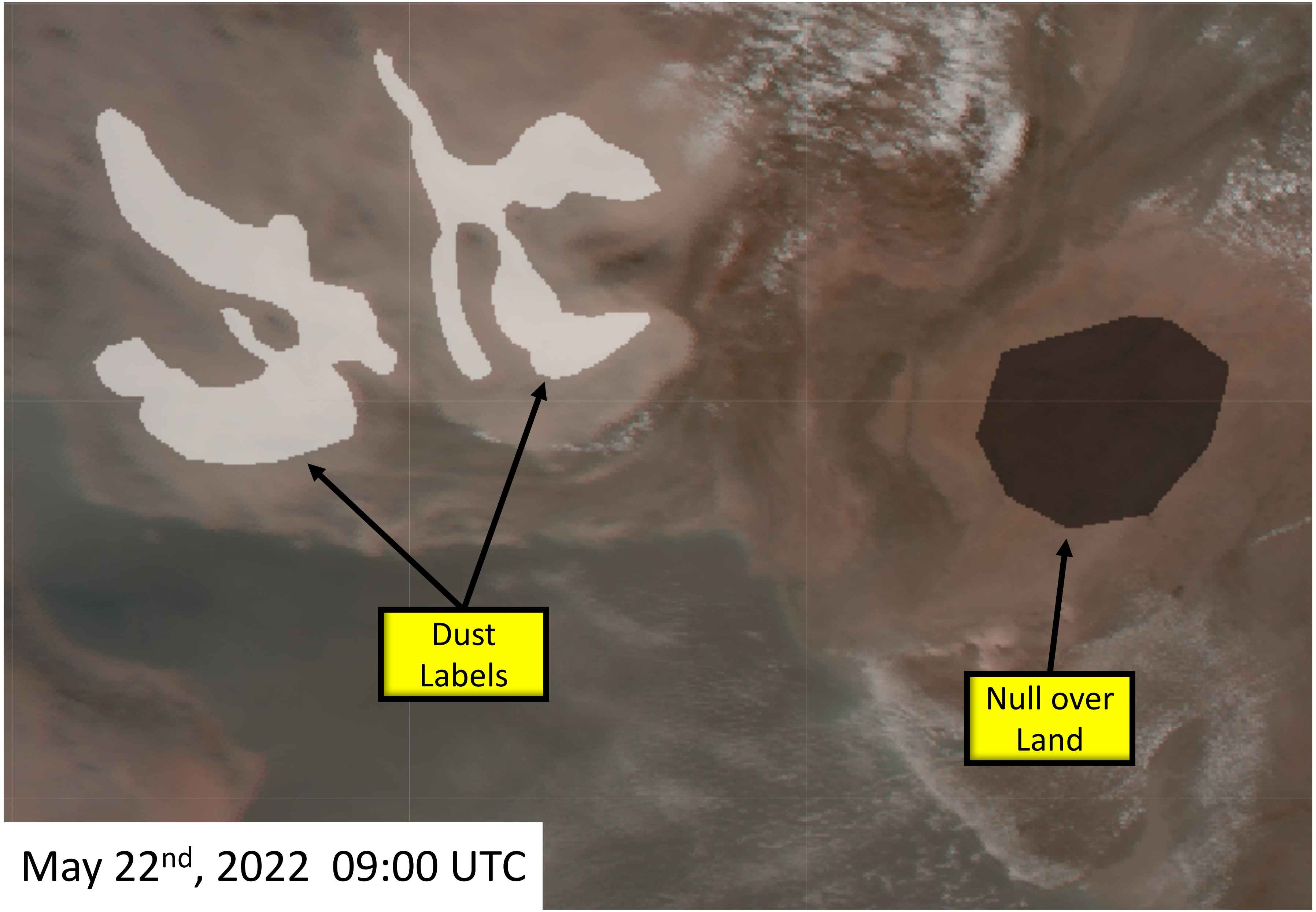


Example Training Data – Dust RGB Labeled



May 22nd, 2022 09:00 UTC

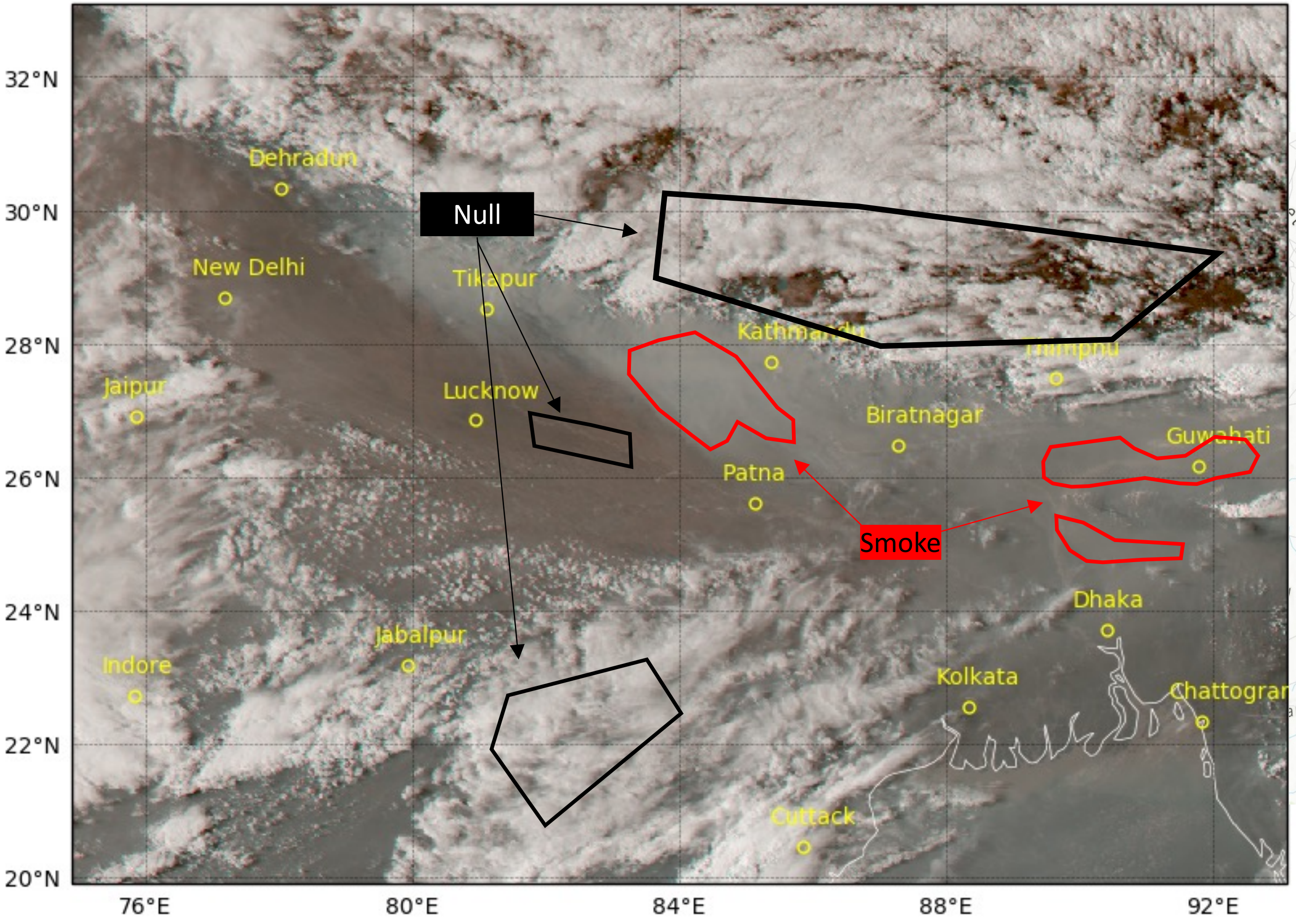
Example Training Data – True Color RGB Labeled

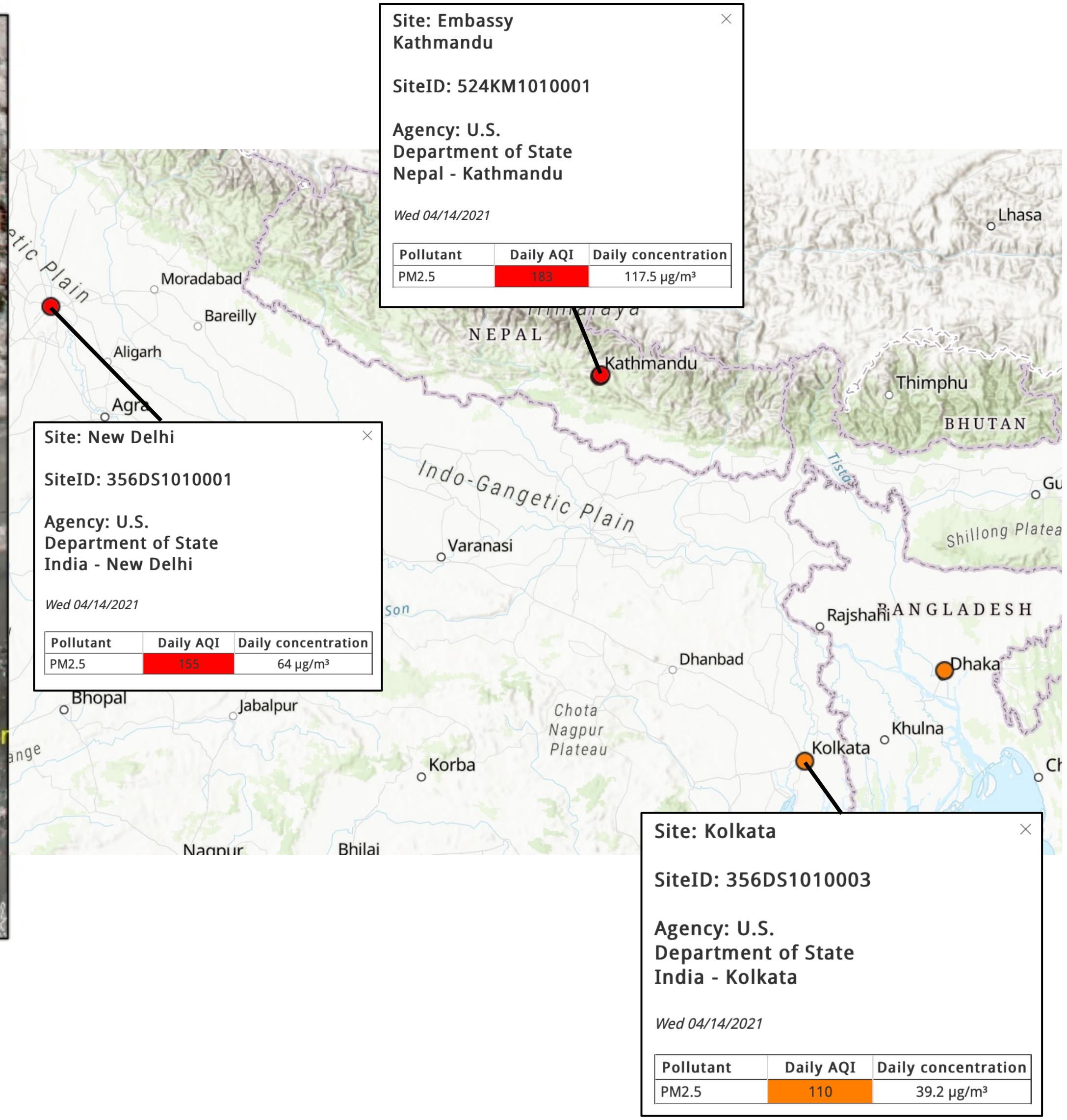


May 22nd, 2022 09:00 UTC

Example Training Data – Smoke Labeled (True Color)

GK2A AMI Truecolor RGB valid 1000 UTC 14 Apr 2021





References

Finding Dust at Night Visualization: https://svs.gsfc.nasa.gov/5040/

Scientific Publication: (Berndt et al. 2021)
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2021EA001788

Real-time DustTracker-AI product:
https://weather.ndc.nasa.gov/sport/viewer/?dataset=goeseastdustmodel&product=dust_prob

NASA Earthdata Feature Article: https://www.earthdata.nasa.gov/learn/articles/dust-ml

SPoRT Twitter Posts:
https://twitter.com/NASA_SPoRT/status/1630668959883096067
https://twitter.com/NASA_SPoRT/status/1648348153991634945
https://twitter.com/NASA_SPoRT/status/1631457114605379587
https://twitter.com/NASA_SPoRT/status/1630668959883096067

Dust ML Paper
(Berndt et al. 2021)

HKH Project Webpage

NASA Visualization Studio
Finding Dust at Night

Backup Content (for Q&A)

General Formulae for RGB Imagery

$$Red = \left(\frac{R - R_{min}}{R_{max} - R_{min}}\right)^{1/\gamma}$$

$$Green = \left(\frac{G - G_{min}}{G_{max} - G_{min}}\right)^{1/\gamma}$$

$$Blue = \left(\frac{B - B_{min}}{B_{max} - B_{min}}\right)^{1/\gamma}$$

Where:

- *R, G,* or *B* is the present pixel value brightness temperature
- *min* and *max* are the calibrated thresholds applied to a given channel or channel difference
- $1/\gamma$ is the calibrated power scale to affect the color stretching

Satellite Products for Air Quality Monitoring

Satellite Product	*Red*	*Green*	*Blue*	*Gamma*	*Applications*
Dust RGB	$IR_{12.3} - IR_{10.5}$ (-6.7 to +2.6°C)	$IR_{11.2} - IR_{8.7}$ (-0.5 to +20°C)	$IR_{10.5}$ (-11.95 to +15.55°C)	1.0 (RB) 2.5 (G)	**Dust plume monitoring**
Nighttime Microphysics (night only)	$IR_{12.3} - IR_{10.5}$ (-6.7 to +2.6°C)	$IR_{10.5} - SW_{3.8}$ (-3.1 to +5.2°C)	$IR_{10.5}$ (-29.55 to +19.45°C)	1.0	**Fog, smog, and low-cloud detection**
Truecolor RGB (day only)	$VIS_{0.64}$ (0 to 1.0 refl)	$VIS_{0.51}$ (0 to 1.0 refl)	$VIS_{0.47}$ (0 to 1.0 refl)	2.2	**Land surface, clouds and smoke**
Natural Color Fire RGB (day only)	$SW_{3.8}$ (0 to 60°C)	$VIS_{0.87}$ (0 to 1.0 refl)	$VIS_{0.64}$ (0 to 1.0 refl)	0.4 (R) 1.0 (GB)	**Fire hot spots [and smoke]**
Fire hot spot detection (GEO + LEO)	**AMI channels combined with** land type, sfc temp, and **MODIS/VIIRS** to identify fire locations at **hourly frequency**				**Fires; early warning on smoke hazards**
Hourly Composite AOD	Uses suite of AMI VIS and IR channels to provide high-quality depiction of total-columnar atmospheric aerosols				**air pollution / data assimilation**
Hourly Composite $PM_{2.5}$	Surface $PM_{2.5}$ **derived from hourly AOD and WRF-Chem** with **data assimilation** model output				**Air quality and health**

Dust Tracker-AI Methodology

Short-term Prediction Research and Transition Center

- Overcoming the limitations of night-time dust detection is addressed by using expert analysis and remote sensing principles to develop training data, model inputs, and architecture.

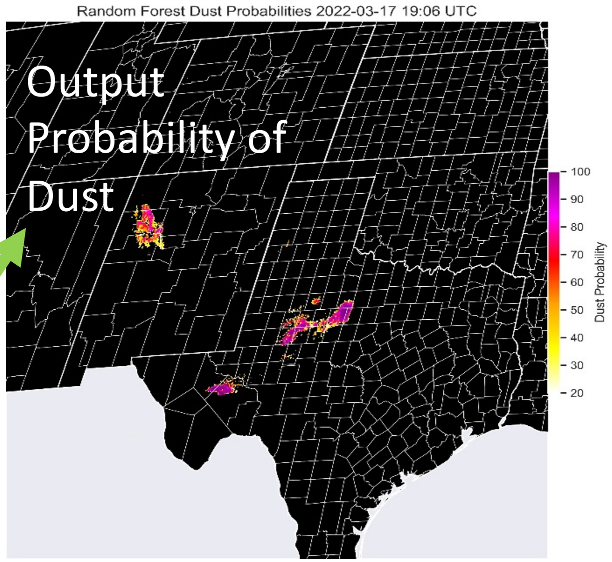
Focus on refining the training for nocturnal dust detection

- Collect night-time training dataset
- Classify false surface and smoke detections

- **GOES-16 ABI imagery in the SW U.S. Jan. 2018 - Jun. 2020**
- **Cases were randomly split for Dust ML model training (60%), testing (20%), and validation (20%)**
- **A <u>total of 28 cases</u> were gathered, which incorporates 83 distinct images a total of 790,921 dust pixels and 37,698,467 null pixels**

<u>Model Inputs</u>
Single channels: 7.3, 10.35, 11.2, 12.3, 13.3
Differences: 12.3-10.35
11.2-8.4
Dust RGB components

Random Forest Model



Evaluate Training, Model Inputs, & Performance w/ Stat Approaches

- Loss functions/Jaccard score
- Confusion matrix
- Permutation Importance
- ROC/AUC
- Plus a few more

Day / Night Dust Tracker AI Model

- Total of 39 training Day & Night cases were gathered,
  - Incorporates 115 distinct images
  - Total of **1,154,064 dust** pixels and **256,932,301 null** pixels
- Assessed the incorporation of visible channels